

C6

## (12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau(43) International Publication Date  
9 August 2001 (09.08.2001)

PCT

(10) International Publication Number  
WO 01/57190 A2(51) International Patent Classification<sup>7</sup>: C12N

(21) International Application Number: PCT/US01/04098

(22) International Filing Date: 5 February 2001 (05.02.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

09/496,914	3 February 2000 (03.02.2000)	US
09/560,875	27 April 2000 (27.04.2000)	US
09/598,075	20 June 2000 (20.06.2000)	US
09/620,325	19 July 2000 (19.07.2000)	US
09/654,936	1 September 2000 (01.09.2000)	US
09/663,561	15 September 2000 (15.09.2000)	US
09/693,325	20 October 2000 (20.10.2000)	US
09/728,422	30 November 2000 (30.11.2000)	US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US	09/496,914 (CIP)
Filed on	3 February 2000 (03.02.2000)
US	09/560,875 (CIP)
Filed on	27 April 2000 (27.04.2000)
US	09/598,075 (CIP)
Filed on	20 June 2000 (20.06.2000)
US	09/620,325 (CIP)
Filed on	19 July 2000 (19.07.2000)
US	09/654,936 (CIP)
Filed on	1 September 2000 (01.09.2000)
US	09/663,561 (CIP)
Filed on	15 September 2000 (15.09.2000)
US	09/693,325 (CIP)
Filed on	20 October 2000 (20.10.2000)
US	09/728,422 (CIP)
Filed on	30 November 2000 (30.11.2000)

(71) Applicant (for all designated States except US): HYSEQ, INC. [US/US]; 670 Almanor Avenue, Sunnyvale, CA 94086 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): TANG, Y., Tom [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). LIU, Chenghua [CN/US]; 1125 Ranchero Way, Apt. #14, San Jose, CA 95117 (US). DRMANAC, Radoje, T. [YU/US]; 850 Greenwich Place, Palo Alto, CA 94303 (US). ASUNDI, Vinod [US/US]; 709 Foster City

Boulevard, Foster City, CA 94404 (US). ZHOU, Ping [CN/US]; 7595 Newcastle Drive, Cupertino, CA 95014 (US). XU, Chongjun [US/US]; 4918 Manitoba Drive, San Jose, CA 95130 (US). CAO, Yicheng [CN/US]; 260 N. Mathilda Avenue, #E6, Sunnyvale, CA 94086 (US). MA, Yunquing [CN/US]; 280 W. California Avenue, #206, Sunnyvale, CA 94086 (US). ZHAO, Qing, A. [CN/US]; 1556 Kooser Road, San Jose, CA 95118 (US). WANG, Dunrui [CN/US]; 932 La Palma Place, Milpitas, CA 95035 (US). WANG, Jian-Rui [CN/US]; 744 Stendahl Lane, Cupertino, CA 95014 (US). ZHANG, Jie [CN/US]; 4930 Poplar Terrace, Campbell, CA 95008 (US). REN, Feiyan [US/US]; 7703 Oak Meadow Court, Cupertino, CA 95014 (US). CHEN, Rui-hong [US/US]; 1031 Flying Fish Street, Foster City, CA 94404 (US). WANG, Zhi, Wei [CN/US]; 836 Alturas Avenue, #B6, Sunnyvale, CA 94085 (US). XUE, Aidong, J. [CN/US]; 1621 S. Mary Avenue, Sunnyvale, CA 94087 (US). YANG, Yonghong [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). WEJHRMAN, Tom [US/US]; CCSR Mol Pharm 3210, 269 W. Campus Drive, Stanford, CA 94305 (US). GOODRICH, Ryle [US/US]; 4896 Sandy Lane, San Jose, CA 95124 (US).

(74) Agent: ELRIFI, Ivor, R.; Mintz, Levin, Cohn, Ferris, Glovsky, and Popeo, P.C., One Financial Center, Boston, MA 02111 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, EU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

## Published:

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

(57) Abstract: The present invention provides novel nucleic acids, novel polypeptide sequences encoded by these nucleic acids and uses thereof.

WO 01/57190 A2

**WHAT IS CLAIMED IS:**

1. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of SEQ ID NO:1-984, 1969-2952, 3937-3942 or 3949-3954, a full length protein coding portion of SEQ ID NO:1-984, 1969-2952, 3937-3942 or 3949-3954, a mature protein coding portion of SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954, an active domain coding portion of SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954, and complementary sequences thereof.
2. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide hybridizes to the polynucleotide of claim 1 under stringent hybridization conditions.
3. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide has greater than about 90% sequence identity with the polynucleotide of claim 1.
4. The polynucleotide of claim 1 wherein said polynucleotide is DNA.
5. An isolated polynucleotide of claim 1 wherein said polynucleotide comprises the complementary sequences.
6. A vector comprising the polynucleotide of claim 1.
7. An expression vector comprising the polynucleotide of claim 1.
8. A host cell genetically engineered to comprise the polynucleotide of claim 1.
9. A host cell genetically engineered to comprise the polynucleotide of claim 1 operatively associated with a regulatory sequence that modulates expression of the polynucleotide in the host cell.
10. An isolated polypeptide, wherein the polypeptide is selected from the group consisting of:
  - (a) a polypeptide encoded by any one of the polynucleotides of claim 1; and
  - (b) a polypeptide encoded by a polynucleotide hybridizing under stringent conditions with any one of SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954.

11. A composition comprising the polypeptide of claim 10 and a carrier.
12. An antibody directed against the polypeptide of claim 10.
13. A method for detecting the polynucleotide of claim 1 in a sample, comprising:
  - a) contacting the sample with a compound that binds to and forms a complex with the polynucleotide of claim 1 for a period sufficient to form the complex; and
  - b) detecting the complex, so that if a complex is detected, the polynucleotide of claim 1 is detected.
14. A method for detecting the polynucleotide of claim 1 in a sample, comprising:
  - a) contacting the sample under stringent hybridization conditions with nucleic acid primers that anneal to the polynucleotide of claim 1 under such conditions;
  - b) amplifying a product comprising at least a portion of the polynucleotide of claim 1; and
  - c) detecting said product and thereby the polynucleotide of claim 1 in the sample.
15. The method of claim 14, wherein the polynucleotide is an RNA molecule and the method further comprises reverse transcribing an annealed RNA molecule into a cDNA polynucleotide.
16. A method for detecting the polypeptide of claim 10 in a sample, comprising:
  - a) contacting the sample with a compound that binds to and forms a complex with the polypeptide under conditions and for a period sufficient to form the complex; and
  - b) detecting formation of the complex, so that if a complex formation is detected, the polypeptide of claim 10 is detected.
17. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:
  - a) contacting the compound with the polypeptide of claim 10 under conditions sufficient to form a polypeptide/compound complex; and
  - b) detecting the complex, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.

18. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:

- a) contacting the compound with the polypeptide of claim 10, in a cell, under conditions sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence in the cell; and
- b) detecting the complex by detecting reporter gene sequence expression, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.

19. A method of producing the polypeptide of claim 10, comprising,

- a) culturing a host cell comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954, a mature protein coding portion of SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954, an active domain coding portion of SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954, complementary sequences thereof and a polynucleotide sequence hybridizing under stringent conditions to SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954, under conditions sufficient to express the polypeptide in said cell; and

- b) isolating the polypeptide from the cell culture or cells of step (a).

20. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of any one of the polypeptides SEQ ID NO: 985-1968, 2953-3936, 3943-3948 or 3955-3960, the mature protein portion thereof, or the active domain thereof.

21. The polypeptide of claim 20 wherein the polypeptide is provided on a polypeptide array.

22. A collection of polynucleotides, wherein the collection comprising the sequence information of at least one of SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954.

23. The collection of claim 22, wherein the collection is provided on a nucleic acid array.

24. The collection of claim 23, wherein the array detects full-matches to any one of the polynucleotides in the collection.

25. The collection of claim 23, wherein the array detects mismatches to any one of the polynucleotides in the collection.

26. The collection of claim 22, wherein the collection is provided in a computer-readable format.
27. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.
28. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising an antibody that specifically binds to a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.



SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
115	AF110399	Homo sapiens	elongation factor Ts	1666	100
116	AF210317	Homo sapiens	facilitative glucose transporter family member GLUT9	2052	99
117	Y73328	Homo sapiens	HTRM clone 082843 protein sequence.	931	100
118	X04085	Homo sapiens	catalase	2846	100
119	AF147717	Homo sapiens	ubiquitin C-terminal hydrolase UCH37	1695	100
120	X73882	Homo sapiens	microtubule associated protein	3801	99
121	AC004882	Homo sapiens	similar to CAA16821 (PID:g3255952)	3223	100
122	M93311	Homo sapiens	metallothionein-III	421	100
123	G03827	Homo sapiens	Human secreted protein, SEQ ID NO: 7908.	557	94
124	G03827	Homo sapiens	Human secreted protein, SEQ ID NO: 7908.	222	53
125	AF232009	Homo sapiens	peroxisomal trans 2-enoyl CoA reductase	1565	99
126	AB004906	Ipomoea purpurea	transposase	146	20
127	M60165	Homo sapiens	guanine nucleotide-binding regulatory protein 2	1832	99
128	Y10319	Homo sapiens	carnitine carrier	1592	100
129	U75467	Drosophila melanogaster	Atu	937	36
130	Z21507	Homo sapiens	human elongation factor-1-delta	494	87
131	Z21507	Homo sapiens	human elongation factor-1-delta	938	100
132	Y58633	Homo sapiens	Protein regulating gene expression PRGE-26.	6745	100
133	Y58633	Homo sapiens	Protein regulating gene expression PRGE-26.	4818	95
134	M13692	Homo sapiens	alpha-1 acid glycoprotein precursor	1064	99
135	U72970	Sus scrofa	calcium/calmodulin-dependent protein kinase II isoform gamma-B	2723	99
136	G03213	Homo sapiens	Human secreted protein, SEQ ID NO: 7294.	450	100
137	AC005102	Homo sapiens	small inducible cytokine subfamily A member 24	627	99
138	AF155648	Homo sapiens	putative zinc finger protein	5855	92
139	AF144638	Homo sapiens	sphingosine-1-phosphate lyase	2977	100
140	AF152318	Homo sapiens	protocadherin gamma A1	4778	100
141	B08517	Homo sapiens	Amino acid sequence of a beta-tubulin antigen.	5841	100
142	X56667	Homo sapiens	calretinin	1410	99
143	X92763	Homo sapiens	tafazzins	1605	100
144	Y95293	Homo sapiens	Human GEF containing NEK-like kinase substrate sGNK.	4092	99
145	AF226046	Homo sapiens	GK003	1198	100
146	M22877	Homo sapiens	cytochrome c	554	98
147	AJ272212	Homo sapiens	protein serine kinase	2196	100
148	AB026491	Homo sapiens	PICK1	2114	98
149	AB018580	Homo sapiens	hluPGFS	1699	100
150	X91868	Homo sapiens	six1	1509	100
151	AF266505	Mus musculus	pseudouridine synthase 3	2135	84
152	U29170	Drosophila melanogaster	ANON-23D-	883	43
153	G04075	Homo sapiens	Human secreted protein, SEQ ID NO: 8156.	567	99
154	AY009128	Homo sapiens	ISCU2	138	100



SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
		BINDING (CREB) PROTEIN SIGNATURE	
124	PR00041	CAMP RESPONSE ELEMENT BINDING (CREB) PROTEIN SIGNATURE	PR00041D 7.95 2.906e-09 24-41
125	BL00061	Short-chain dehydrogenases/reductases family proteins.	BL00061C 7.86 3.250e-10 212-222
126	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 6.400e-25 251-290
127	PR00318	ALPHA G-PROTEIN (TRANSDUCIN) SIGNATURE	PR00318D 16.28 1.900e-34 219-248 PR00318B 14.79 3.455e-27 168-191 PR00318C 12.09 7.000e-23 197-215 PR00318A 7.84 1.600e-19 35-51 PR00318E 7.23 2.500e-12 265-275
128	PR00927	ADENINE NUCLEOTIDE TRANSLOCATOR 1 SIGNATURE	PR00927E 14.93 9.743e-10 67-89 PR00927B 14.66 4.575e-09 69-91
130	BL00824	Elongation factor 1 beta/beta'/delta chain proteins.	BL00824B 9.21 7.750e-22 133-153
131	BL00824	Elongation factor 1 beta/beta'/delta chain proteins.	BL00824C 14.58 1.000e-40 166-204 BL00824D 14.04 1.621e-38 204-239 BL00824B 9.21 7.750e-22 133-153 BL00824E 12.49 1.000e-19 247-263
132	PR00209	ALPHA/BETA GLIADIN FAMILY SIGNATURE	PR00209B 4.88 9.222e-13 1209-1228
133	PR00209	ALPHA/BETA GLIADIN FAMILY SIGNATURE	PR00209B 4.88 9.222e-13 1168-1187
134	PR00708	ALPHA-1-ACID GLYCOPROTEIN SIGNATURE	PR00708D 14.67 1.000e-27 141-168 PR00708C 11.77 1.643e-25 98-120 PR00708B 15.15 2.174e-24 73-95 PR00708E 13.33 1.600e-21 189-207 PR00708A 14.40 2.636e-21 51-70
135	PR00109	TYROSINE KINASE CATALYTIC DOMAIN SIGNATURE	PR00109B 12.27 8.468e-13 126-145
136	PF00023	Ank repeat proteins.	PF00023A 16.03 3.250e-10 201-217
137	BL00471	Small cytokines (intercrine/chemokine) C-x-C subfamily signat.	BL00471 23.92 7.480e-10 42-90
140	PR00205	CADHERIN SIGNATURE	PR00205B 11.39 5.582e-10 328-346 PR00205B 11.39 9.018e-10 543-561
141	BL00412	Neuromodulin (GAP-43) proteins.	BL00412D 16.54 7.704e-09 976-1027
143	PR00979	TAFAZZIN SIGNATURE	PR00979E 10.83 5.950e-26 192-214 PR00979A 11.91 8.773e-25 63-83 PR00979C 12.16 6.400e-19 108-124 PR00979D 12.38 7.955e-19 170-185 PR00979F 10.14 3.382e-15 230-244 PR00979B 15.59 5.636e-15 94-106
145	DM00686	kw REPLICATION REP 28K 17.7K.	DM00686C 14.14 7.720e-09 111-131
146	PR00604	CLASS IA AND IB CYTOCHROME C SIGNATURE	PR00604D 15.86 1.000e-17 87-104 PR00604B 12.73 9.591e-16 57-73 PR00604C 10.21 8.200e-12 73-84 PR00604E 10.13 1.000e-11 106-117 PR00604A 11.13 8.800e-



SEQ ID NO:	PFAM NAME	DESCRIPTION	p-value	PFAM SCORE
112	HSP20	Hsp20/alpha crystallin family	2.6e-20	77.7
115	EF_TS	Elongation factor TS	3.8e-63	221.1
116	sugar_tr	Sugar (and other) transporter	4e-63	223.1
118	catalase	Catalase	0	1158.9
119	UCH	Ubiquitin carboxyl-terminal hydrolase, famil	1e-10	24.4
122	metalthio	Metallothionein	2.8e-25	97.4
125	adh_short	short chain dehydrogenase	1.6e-45	164.6
126	KRAB	KRAB box	7.9e-25	95.9
127	G-alpha	G-protein alpha subunit	1e-249	843.0
128	mito_carr	Mitochondrial carrier proteins	2e-65	227.2
131	EF1BD	EF-1 guanine nucleotide exchange domain	4.9e-53	189.6
132	GYF	GYF domain	4.9e-28	106.6
133	GYF	GYF domain	4.9e-28	106.6
134	lipocalin	Lipocalin / cytosolic fatty-acid binding pr	2.1e-33	119.1
135	pkinase	Eukaryotic protein kinase domain	3.3e-86	299.8
136	ank	Ank repeat	2.2e-29	111.1
137	IL8	Small cytokines (intecrine/chemokine), inter	3.1e-18	65.2
139	pyridoxal_deC	Pyridoxal-dependent decarboxylase conse	0.00011	19.0
140	cadherin	Cadherin domain	1.3e-88	307.8
142	efhand	EF hand	5.7e-33	123.0
143	Acytransferase	Acytransferase	2e-29	111.2
146	cytochrome_c	Cytochrome c	1.7e-33	124.7
147	pkinase	Eukaryotic protein kinase domain	2.3e-86	300.3
148	PDZ	PDZ domain (Also known as DHR or GLGF).	1.7e-09	45.0
149	aldo_ket_red	Aldo/keto reductase family	7.4e-189	640.8
150	homeobox	Homeobox domain	3.2e-08	38.7
151	PseudoU_synth_1	tRNA pseudouridine synthase	4.7e-57	203.0
152	abhydrolase	alpha/beta hydrolase fold	1.7e-31	118.0
153	PDZ	PDZ domain (Also known as DHR or GLGF).	1.1e-09	45.6
156	PHD	PHD-finger	7.6e-15	62.8
157	fn3	Fibronectin type III domain	0.015	21.9
158	homeobox	Homeobox domain	2.7e-27	104.1
160	PWI	PWI domain	3.9e-24	93.6
162	DnaJ	DnaJ domain	2e-06	34.8
164	Cbl_N	CBL proto-oncogene N-terminal domain	8e-117	401.5
166	metalthio	Metallothionein	3.1e-26	100.6
167	LRR	Leucine Rich Repeat	0.00069	26.3
169	fibrinogen_C	Fibrinogen beta and gamma chains, C-term	5.3e-180	611.4
170	fibrinogen_C	Fibrinogen beta and gamma chains, C-term	5.3e-180	611.4
171	fibrinogen_C	Fibrinogen beta and gamma chains, C-term	1e-149	510.8
173	homeobox	Homeobox domain	1.5e-29	111.6
174	FYVE	FYVE zinc finger	7.4e-28	103.8
175	GRIP	GRIP domain	3.9e-08	40.5
182	pkinase	Eukaryotic protein kinase domain	3.4e-71	250.0
185	CAP_GLY	CAP-Gly domain	5.6e-51	182.8
186	TBC	TBC domain	2.2e-50	180.8
187	TBC	TBC domain	2.2e-50	180.8



114	1098	2082	3066	787CIP2_116	8004
115	1099	2083	3067	787CIP2_117	8007
116	1100	2084	3068	787CIP2_118	8008
117	1101	2085	3069	787CIP2_119	8009
118	1102	2086	3070	787CIP2_120	8013
119	1103	2087	3071	787CIP2_121	8017
120	1104	2088	3072	787CIP2_122	8018
121	1105	2089	3073	787CIP2_123	8021
122	1106	2090	3074	787CIP2_124	8022
123	1107	2091	3075	787CIP2_125	8023
124	1108	2092	3076	787CIP2_126	8023
125	1109	2093	3077	787CIP2_127	8024
126	1110	2094	3078	787CIP2_128	8026
127	1111	2095	3079	787CIP2_129	8028
128	1112	2096	3080	787CIP2_130	8036
129	1113	2097	3081	787CIP2_131	8038
130	1114	2098	3082	787CIP2_132	8045
131	1115	2099	3083	787CIP2_133	8045
132	1116	2100	3084	787CIP2_134	8048
133	1117	2101	3085	787CIP2_135	8048
134	1118	2102	3086	787CIP2_136	8052
135	1119	2103	3087	787CIP2_137	8053
136	1120	2104	3088	787CIP2_138	8055
137	1121	2105	3089	787CIP2_139	8059
138	1122	2106	3090	787CIP2_140	8061
139	1123	2107	3091	787CIP2_141	8062
140	1124	2108	3092	787CIP2_142	8063
141	1125	2109	3093	787CIP2_143	8064
142	1126	2110	3094	787CIP2_144	8065
143	1127	2111	3095	787CIP2_145	8068
144	1128	2112	3096	787CIP2_146	8069
145	1129	2113	3097	787CIP2_147	8070
146	1130	2114	3098	787CIP2_148	8074
147	1131	2115	3099	787CIP2_149	8076
148	1132	2116	3100	787CIP2_150	8077
149	1133	2117	3101	787CIP2_151	8078
150	1134	2118	3102	787CIP2_152	8079
151	1135	2119	3103	787CIP2_153	8087
152	1136	2120	3104	787CIP2_154	8091
153	1137	2121	3105	787CIP2_155	8100
154	1138	2122	3106	787CIP2_156	8105
155	1139	2123	3107	787CIP2_157	8106
156	1140	2124	3108	787CIP2_158	8108
157	1141	2125	3109	787CIP2_159	8109
158	1142	2126	3110	787CIP2_160	8110
159	1143	2127	3111	787CIP2_161	8112
160	1144	2128	3112	787CIP2_162	8116
161	1145	2129	3113	787CIP2_163	8118
162	1146	2130	3114	787CIP2_164	8124
163	1147	2131	3115	787CIP2_165	8125
164	1148	2132	3116	787CIP2_166	8127
165	1149	2133	3117	787CIP2_167	8132
166	1150	2134	3118	787CIP2_168	8135
167	1151	2135	3119	787CIP2_169	8137
168	1152	2136	3120	787CIP2_170	8139
169	1153	2137	3121	787CIP2_171	8140
170	1154	2138	3122	787CIP2_172	8140
171	1155	2139	3123	787CIP2_173	8140
172	1156	2140	3124	787CIP2_174	8141
173	1157	2141	3125	787CIP2_175	8147



SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion,  =possible nucleotide insertion)
				PNRARNNTHSNLHTSIGNSVWGSINTGPPNQWA SDLVSSIWSNADTKNSNMGFWDAAVKEVGPRN STNKNKNNASLSKSVGVSNRQNKVEEEKLLK LFQGVNKAQDGFTQWCEQMLHALNTANNLDVP TFVSFLKEVESPYEVHDYIRAYLGDTSEAKEFAK QFLERRAKQKANQQRQQQLPQQQQQPPQQPP QQPQQQDSDVWGMNHSTLHSVFQTNQSNNQQSN FEAVQSGKKKKQKMKVRADPSLLGFSVNASSER LNMGEIETLDDY
3086	A	675	1334	LHPAATSTA WLHVPPGLSMALSWVLTVLSSLPL LEAQIPLCANLVPVPITNATLDRITGKWFYIASAF RNEEYNKSVQEIQATFFYFTPNTKTEDTIFLREYQT RQDQCIYNTTYLNVQRENGTISRYVGGQEFAH LLILRDTKTYMLAFDVNDEKNWGLSVYADKPET TKEQLGEFYEALDCLRIPKSDVVYTDWKKDKCE PLEKQHEKERKQEEGES
3087	A	1	1575	CTPVARSMAATTCTRFTDDYQLFEELGKGAFS VVRRCVKKTSTQEYAAKIINTKKLSARDHQKLE REARICRLLKHPNIVRLHDSISEEGFHYLVFDLVT GGELFEDIVAREYYSEADASHCIHQILESVDNHIHQ HDIVHRDLKPENLLLASKCKGAAVKLADFGGLAIE VQGEQQAWFGFAGTPGYLSPEVLRKDGYGKPWD IWACGVILYILLVGYPFWDEDQHKL YQQIKAG AYDFPSPEWDTVTPEAKNLINQMLTINPAKRITA DQALKHPWVCQRSTVASMMHRQETVECLRKFN ARRKLKGAILTTMLVSRNFSAAKSLLNKKSDGG VKPQSNNKNSLVSPAQEPAQLQTAMEPQTTVVH NATDGIGSTESCNTTEDEDLKVRKQEIIKITEQ LIEAINNGDFEA YTKICDPGLTSFEPEALGNL VEG MDFHKFYFENLLSKNSKPIHTTILNPHVHVIGED AACIA YIRLTQYIDGQGRPRTSQSEETRVWHRD GKWLNVHYHCSGAPAQLQ
3088	A	12	1039	SSVAEFPERVQLSQPNWNFGAGGAWSLDFAE QLKWSAELARLGESIMDGKQGGMDGSKPAGPR DFPGIRLLSNPLMGDAVSDWSPMHEAAIHGHQL SLRNLISQGWAVNITADHVSPLHEACLGGHLSC VKILLKHGAQVNGVTADWHTPLFNACVSGSWD CVNLLLQHGASVQPESDLASPIHEAARRGHVEC VNSLIAYGGNIDHKISHLGTPLYLACENQQRACV KKLLESGADVNGKGQDPLHA VARTASEELAC LLMDFGADTQAKNAEGKRPVELVPPESPLAQLF LEREGPPSLMQLCRLRIRKCFGIQQHHKITKLVLP EDLKQFLLHL
3089	A	73	432	DMAGLMTIVTSLFLGVCAHHIPTGSVVLPSGCC MFFVSKRIPENRVVSYQLSSRSTCLKAGVIFTKK GQQFCGDPKQEWWQRYMKNLDAKQKKASPRA RAAVAKGPVQRYPGNQTC
3090	A	4627	611	LMEAGGGGGALPAGVETMVLTLGESWPVLVGR RFLSLSAADGSDGSHDSWDVERVAEWPWLSGTI RAVSHTDVTKKDLKVCVEFDGESWRKRRWIEV YSLLRRAFLVEHNLVLAERKSPEISERIVQWPAIT YKPLLDKAGLGSITSVRFLGDQQQRVFLSKDLLKP IQDVNSLRLSLTDNQIVSKEFQALIVKHLDESHLL KGDKNLVGSEVKIYSLDPSTQWFSATVVNGNPA SKTLQVNCEEIPALKIVDPSLIHVEVVHDNLVTC



Pages 485 to 6221 of this application contain amino acid sequence listings. They can be obtained at the address given below.

Les pages 485 to 6221 de cette demande contiennent des listages des séquences d'acides aminés. Elles peuvent être obtenues à l'adresse indiquée ci-dessous.

World Intellectual Property Organization  
34, chemin des Colombettes  
CH-1211 Genève 20

